

# ERGODIC THEORY AND DIOPHANTINE APPROXIMATION FOR TRANSLATION SURFACES AND LINEAR FORMS

JAYADEV ATHREYA, ANDREW PARRISH, AND JIMMY TSENG

**ABSTRACT.** We derive results on the distribution of directions of saddle connections on translation surfaces, using only the Birkhoff ergodic theorem applied to the geodesic flow on the moduli space of translation surfaces. Our techniques, together with an approximation argument, also give a simple proof of a weak version of a classical theorem in multi-dimensional Diophantine approximation due to W. Schmidt [23]. The approximation argument allows us to deduce the Birkhoff genericity of almost all lattices in a certain submanifold of the space of unimodular lattices from the Birkhoff genericity of almost all lattices in the whole space and similarly for the space of affine unimodular lattices.

## 1. INTRODUCTION

In this paper, we study translation surfaces, linear and affine forms, toral translations, and unimodular lattices of  $\mathbb{R}^d$  using a few simple tools from ergodic theory. This provides a unified and simplified viewpoint and allows us to derive new results and explain classical ones.

**1.1. Translation Surfaces.** A *translation surface* is a pair  $(M, \omega)$ , where  $M$  is a Riemann surface and  $\omega$  a holomorphic 1-form. We refer the reader to Zorich [31] for an excellent survey on translation surfaces. Often, we will use simply  $\omega$  to refer to a translation surface. A *saddle connection* on a translation surface  $\omega$  is a geodesic  $\gamma$  (in the flat metric induced by  $\omega$ ) connecting two zeros of  $\omega$  (with none in its interior). Moreover, to each saddle connection  $\gamma$ , one can associate a holonomy vector  $\mathbf{v}_\gamma = \int_\gamma \omega \in \mathbb{C}$ .

The set of holonomy vectors  $\Lambda_\omega := \Lambda_{sc}(\omega)$  is a discrete subset of  $\mathbb{C} \cong \mathbb{R}^2$ . Saddle connections arise naturally as special trajectories for billiards in rational-angled polygons; see, for example, [21]. As an example, given a unimodular lattice  $\Lambda \subset \mathbb{C}$ , the associated flat torus  $\mathbb{C}/\Lambda$  is a translation surface, and, ‘marking’ the point 0 as a zero of  $\omega$ , the set of saddle connections corresponds to the set of *primitive* vectors in  $\Lambda$ , that is, the set of vectors in  $\Lambda$  which are not nontrivial multiples of other vectors in  $\Lambda$ .

Understanding the geometry of the set  $\Lambda_\omega$  has been one of the central themes in the study of translation surfaces. Masur [20], Veech [29] and Eskin-Masur [9] proved seminal counting results, showing, respectively, that

$$N(\omega, R) := \#(\Lambda_\omega \cap B(0, R))$$

has upper and lower quadratic upper bounds, quadratic asymptotics on average, and quadratic asymptotics for almost every translation surface. Recently, Eskin-Mirzakhani-Mohammadi [10] showed that weak quadratic asymptotics hold for *every* point.

The results of Eskin-Masur [9] also imply that the directions of saddle connection vectors equidistribute on  $S^1$  for almost every surface. The fine-scale distribution has been studied in terms of the gap distribution by Athreya-Chaika [1], Athreya-Chaika-Lelievre [2], and Uyanik-Work [27]. In this paper, we study a counting problem related to the fine-scale distribution properties of  $\Lambda_\omega$ , by counting the saddle connections which approximate the vertical to a prescribed degree. Given a translation surface  $\omega$ ,  $T > 1$ , and  $b > 0$ , let

$$R_{b,T}(\omega) = \# \{(x, y) \in \Lambda_\omega : |xy| \leq b \text{ and } 1 \leq y < T\}$$

denote the function which counts the number of saddle connections in a thinning (hyperbolic) region around the  $y$ -axis.

J.S.A. partially supported by NSF grant DMS 1069153, and NSF grants DMS 1107452, 1107263, 1107367 “RNMS: GEometric structures And Representation varieties” (the GEAR Network), and NSF CAREER grant DMS 1351853.

J.T. acknowledges the research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 291147.

**Theorem 1.1.** *Let  $\mu$  be an ergodic  $\mathrm{SL}(2, \mathbb{R})$ -invariant measure on a stratum  $\mathcal{H}$  of translation surfaces. There is a constant  $C = C(\mu)$  (known as the Siegel-Veech constant of  $\mu$ ) so that for  $\mu$ -a.e.  $\omega$ , we have*

$$\lim_{T \rightarrow \infty} \frac{R_{b,T}(\omega)}{2bC \log T} = 1.$$

In particular, there is a natural absolutely continuous invariant measure  $\mu_{MV}$  on each stratum  $\mathcal{H}$  known as the Masur-Veech measure (for which the action of  $\mathrm{SL}(2, \mathbb{R})$  was shown to be ergodic independently by Masur [19] and Veech [28]), and this theorem applies to  $\mu_{MV}$ -a.e.  $\omega \in \mathcal{H}$ .

**1.1.1. Ergodic Theory, Counting, and Approximation.** The proof of Theorem 1.1 relies on ergodic theory on the moduli space of translation surfaces. The idea to use ergodic theory applied to counting problems in this context comes from work of Veech [29] and Eskin-Masur [9], and was inspired by the work on the Quantitative Oppenheim conjecture by Eskin-Margulis-Mozes [8], which used ergodic theory on homogeneous spaces. These results rely on sophisticated equidistribution theorems, and a large motivation for this paper is to show how even the original (Birkhoff) ergodic theorem can yield interesting geometric information when applied in these contexts.

Theorem 1.1 gives an analog of (a weak form of) Schmidt's theorem for translation surfaces, highlighting the well-known connections between translation surfaces and Diophantine approximation. Our techniques in fact give new, simple proofs of Schmidt's results, which, along with the overall strategy of our proofs, we describe below. Also, one can apply some of these techniques to the study of the distribution of directions of lattice vectors that arise in Diophantine approximation as done in [3, 4].

**1.2. Diophantine Approximation.** Suppose  $\psi_i(n) : \mathbb{N} \rightarrow \mathbb{R}, 1 \leq i \leq k$  are non-negative functions and that  $\psi(n) = \prod_{i=1}^k \psi_i(n)$  is monotonically decreasing. In 1926, A. Khintchine showed that

*Theorem ([12]).* If  $\sum_{n=1}^{\infty} \psi(n)$  diverges, then there are infinitely many solutions  $(n, p_1, \dots, p_k)$  to the system of inequalities

$$|x_i n - p_i| \leq \psi_i(n),$$

for a.e.  $(x_1, \dots, x_k) =: \mathbf{x} \in \mathbb{R}^k$ . If  $\sum_{n=1}^{\infty} \psi(n)$  converges, then there are at most finitely many solutions for a.e.  $\mathbf{x} \in \mathbb{R}^k$ .

This statement was later refined by W. Schmidt, who showed in [23] that the number of solutions of the system of inequalities,

$$|x_i n - p_i| \leq \psi_i(n),$$

with  $1 \leq n \leq h$  is on the order of  $\sum_{n=1}^h \psi(n)$  while also giving an estimate on the size of the error term.

To illustrate the flexibility of our techniques, we show how to give a simple proof of the divergence case of Schmidt's theorem for the functions,

$$\psi_i(n) := \frac{1}{n^{1/k}},$$

using only the Siegel mean value theorem, the Birkhoff ergodic theorem, and an approximation argument. These functions occur naturally in questions of Diophantine approximation. Given  $\mathbf{x} \in \mathbb{R}^k$ ,  $T > 1$ , let  $N(\mathbf{x}, T)$  denote the number of solutions  $(n, \mathbf{p}) \in \mathbb{Z} \times \mathbb{Z}^k =: \mathbb{Z}^{k+1}$  of the system of inequalities

$$\|n\mathbf{x} - \mathbf{p}\| \leq \frac{1}{n^{1/k}}$$

with  $1 \leq n < T$ . Here and below, let  $\|\cdot\|$  denote the Euclidean norm (this is not crucial for our results, but helps to streamline our statements). Let  $B_k$  denote the volume of the unit  $k$ -ball.

**Theorem 1.2.** *For a.e.  $\mathbf{x} \in \mathbb{R}^k$ ,*

$$\lim_{T \rightarrow \infty} \frac{N(\mathbf{x}, T)}{2B_k \log T} = 1.$$

While this statement is weaker than Schmidt's theorem, the proof given below relies only on the ergodicity of a certain flow in the space of unimodular lattices. This is in contrast to the dynamical approaches for theorems closely related to Schmidt's detailed in [5, 15], where mixing properties and representation theory are used.

1.2.1. *Homogeneous linear forms.* In fact, Schmidt stated his result in terms of linear forms, and our results also apply in this setting. Let  $A \in M_{m \times n}(\mathbb{R})$  be an  $m \times n$  matrix, which we view as a system of  $m$  linear forms in  $n$  variables. A classical Diophantine question is to find approximate integral solutions to the equation  $A\mathbf{y} = \mathbf{x}$ ,  $\mathbf{y} \in \mathbb{R}^n$ ,  $\mathbf{x} \in \mathbb{R}^m$ . In this context, a classical theorem of Dirichlet states that, for every  $A$ , there are infinitely many solutions  $\mathbf{p} \in \mathbb{Z}^m$ ,  $\mathbf{q} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$  to the inequality

$$\|A\mathbf{q} - \mathbf{p}\| \leq \|\mathbf{q}\|^{-\frac{n}{m}}.$$

Let  $b > 0$ ,

$$N(A, b, T) := \# \{(\mathbf{p}, \mathbf{q}) \in \mathbb{Z}^m \times \mathbb{Z}^n : \|A\mathbf{q} - \mathbf{p}\| \leq b\|\mathbf{q}\|^{-\frac{n}{m}}, 1 \leq \|\mathbf{q}\| < T\},$$

$C_k$  denote the surface area of the unit sphere  $\mathbb{S}^{k-1} \subset \mathbb{R}^k$  (e.g.,  $C_1 = 2, C_2 = 2\pi, C_3 = 4\pi$ ), and recall that  $B_k$  denotes the volume of the unit  $k$ -ball and  $\|\cdot\|$  denotes the Euclidean norm. Theorem 1.2 is in fact a special case ( $m = 1, n = k$ ) of

**Theorem 1.3.** *For a.e.  $A \in M_{m \times n}(\mathbb{R})$ ,*

$$\lim_{T \rightarrow \infty} \frac{N(A, b, T)}{b^m B_m C_n \log T} = 1.$$

1.2.2. *Inhomogeneous linear forms.* We can also consider a system of *inhomogeneous linear forms* or, alternatively, *affine forms*: given  $A \in M_{m \times n}(\mathbb{R})$  and  $\mathbf{w} \in \mathbb{R}^m$ , we want approximate integral solutions  $(\mathbf{p}, \mathbf{q})$  to the equation  $A\mathbf{q} = \mathbf{p} + \mathbf{w}$ . For  $b > 0$ , define

$$N(A, \mathbf{w}, b, T) := \# \{(\mathbf{p}, \mathbf{q}) \in \mathbb{Z}^m \times \mathbb{Z}^n : \|A\mathbf{q} - \mathbf{p} - \mathbf{w}\| \leq b\|\mathbf{q}\|^{-\frac{n}{m}}, 1 \leq \|\mathbf{q}\| < T\}.$$

**Theorem 1.4.** *For a.e.  $(A, \mathbf{w}) \in M_{m \times n}(\mathbb{R}) \times \mathbb{R}^m$ ,*

$$\lim_{T \rightarrow \infty} \frac{N(A, \mathbf{w}, b, T)}{b^m B_m C_n \log T} = 1.$$

1.3. **Lattices.** Theorems 1.2, 1.3, and 1.4 are consequences of more general theorems in the geometry of numbers. Let  $m, n \geq 1$ ,  $d = m + n$ , and  $\Lambda \subset \mathbb{R}^d$  be a unimodular lattice, that is, a discrete subgroup of covolume 1 (see Section 2). Viewing

$$\mathbb{R}^d := \mathbb{R}^m \times \mathbb{R}^n,$$

we write elements as  $\mathbf{v} = (\mathbf{x}, \mathbf{y})$ ,  $\mathbf{x} \in \mathbb{R}^m$ ,  $\mathbf{y} \in \mathbb{R}^n$ . By an abuse of notation, we think of all of these as *column* vectors. Given  $b > 0$ , define

$$R_{b,T}(\Lambda) := \# \{(\mathbf{x}, \mathbf{y}) \in \Lambda : \|\mathbf{x}\|^m \|\mathbf{y}\|^n \leq b, 1 \leq \|\mathbf{y}\| < T\}.$$

**Theorem 1.5.** *For a.e. unimodular lattice  $\Lambda \subset \mathbb{R}^d$  (with respect to the Haar measure on the space of unimodular lattices), we have that*

$$\lim_{T \rightarrow \infty} \frac{R_{b,T}(\Lambda)}{b B_m C_n \log T} = 1.$$

We also record a statement on affine unimodular lattices: an *affine unimodular lattice*  $\Lambda + \mathbf{v}$  in  $\mathbb{R}^d$  is a translate of a unimodular lattice  $\Lambda \subset \mathbb{R}^d$  by a vector  $\mathbf{v} \in \mathbb{R}^d / \Lambda$ . We define  $R_{b,T}(\Lambda + \mathbf{v})$  as above.

**Theorem 1.6.** *For a.e. affine unimodular lattice  $\Lambda + \mathbf{v} \subset \mathbb{R}^d$  (with respect to the Haar measure on the space of affine unimodular lattices), we have that*

$$\lim_{T \rightarrow \infty} \frac{R_{b,T}(\Lambda + \mathbf{v})}{b B_m C_n \log T} = 1.$$

A direct computation shows that  $b B_m C_n \log T$  is the  $d$ -dimensional volume of the region

$$\mathcal{R}_{b,T} := \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^d : \|\mathbf{x}\|^m \|\mathbf{y}\|^n \leq b, 1 \leq \|\mathbf{y}\| < T\}$$

where we are counting lattice points.

**1.4. Toral Translations.** Our lattice results can also be interpreted in terms of shrinking target properties (or logarithm laws) for toral translations. Fix  $m \geq 1$ , and  $\alpha \in \mathbb{T}^m := \mathbb{R}^m / \mathbb{Z}^m$ . We consider the dynamical system generated by translation by  $\alpha$ , that is, the map  $T_\alpha : \mathbb{T}^m \rightarrow \mathbb{T}^m$ , where

$$T_\alpha \mathbf{x} = \alpha + \mathbf{x}.$$

Let  $\|\mathbf{x}\|_{\mathbb{Z}} = \min\{\|\mathbf{x} - \mathbf{p}\| : \mathbf{p} \in \mathbb{Z}^m\}$ . For  $b > 0$ , we define

$$S_{b,N}(\alpha) := \#\left\{1 \leq n \leq N : \|n\alpha\|_{\mathbb{Z}} < bn^{-1/m}\right\} = \#\left\{1 \leq n \leq N : T_\alpha^n(\mathbf{0}) \in B(\mathbf{0}, bn^{-1/m})\right\}.$$

where  $B(\mathbf{0}, r)$  denotes a  $\|\cdot\|_{\mathbb{Z}}$ -ball in  $\mathbb{T}^m$  of radius  $r > 0$ . Similarly, given  $\mathbf{v} \in \mathbb{T}^m$ , we define

$$S_{b,N}(\alpha, \mathbf{v}) := \#\left\{1 \leq n \leq N : \|n\alpha - \mathbf{v}\|_{\mathbb{Z}} < bn^{-1/m}\right\} = \#\left\{1 \leq n \leq N : T_\alpha^n(\mathbf{0}) \in B(\mathbf{v}, bn^{-1/m})\right\}.$$

As above, let  $B_m$  be the volume of the unit  $\|\cdot\|$ -ball in  $\mathbb{R}^m$ .

**Corollary 1.7.** *For a.e.  $\alpha$ ,*

$$\lim_{N \rightarrow \infty} \frac{S_{b,N}(\alpha)}{b^m B_m \log N} = 1$$

*and, for a.e.  $\alpha$  and a.e.  $\mathbf{x} \in \mathbb{T}^m$ ,*

$$\lim_{N \rightarrow \infty} \frac{S_{b,N}(\alpha, \mathbf{x})}{b^m B_m \log N} = 1.$$

*Proof.* The first equality follows from Theorem 1.2 and the second from the corresponding statement for Theorem 1.4.  $\square$

Corollary 1.7 can be regarded as a strengthening of the logarithm law for toral translations (see [25, Eqns. (5.1) and (5.2)]). For  $m = 1$ , there is, in addition, a logarithm law for circle rotations for any irrational  $\alpha$  [25, Corollary 1.7], which follows from a certain shrinking target property (see [25, Section 1.3] for more details).

**1.5. Badly approximable forms and bounded geodesics.** Finally, we note that our results for forms, lattices, toral translations, and translation surfaces for the stratum of the flat torus cannot be improved from *almost every* to *every* because of the existence of badly approximable systems of affine forms [7, Theorems 1.1 and 1.4] (see also [13] and [26]). For general strata, the existence of bounded geodesics [14, Theorem 1.2] (see also [6, Theorem 1.3]) shows such improvement cannot occur.

**1.6. Strategy of Proof.** There is a common strategy of proof for Theorems 1.1, 1.5, and 1.6. Namely, we express the quantities  $R_T$  as Birkhoff averages of an appropriate ergodic transformation on a moduli space, and apply the Birkhoff ergodic theorem. The limiting integrals on moduli space can be computed by an application of a Siegel (or Siegel-Veech) formula, allowing us to reduce the problem to a volume computation on Euclidean space. Our other results are applications of these theorems and, in particular, for Theorems 1.2 and 1.3, we parametrize  $\mathrm{SL}(d, \mathbb{R})$  in a convenient way and use an approximation argument to reduce our results about unimodular lattices (Theorems 1.5 and 1.6) to number theory.

## 2. THE SPACE OF UNIMODULAR LATTICES

Given a unimodular lattice  $\Lambda \subset \mathbb{R}^d$ , we can write  $\Lambda = g\mathbb{Z}^d$ , where  $g \in \mathrm{SL}(d, \mathbb{R})$  is well-defined up to multiplication on the right by  $\mathrm{SL}(d, \mathbb{Z})$ . That is, we can identify the space of unimodular lattices with the homogeneous space  $X_d := \mathrm{SL}(d, \mathbb{R}) / \mathrm{SL}(d, \mathbb{Z})$ , which is endowed with a measure that comes from the Haar measure  $\mu = \mu_d$  on  $\mathrm{SL}(d, \mathbb{R})$  and is normalized so that  $\mu_d(X_d) = 1$ .

**2.1. Mean value formulas.** A key ingredient of our proof is the computation of the average number (with respect to  $\mu_d$ ) of lattice points in a given subset of  $\mathbb{R}^d$ . This is known as the Siegel mean value theorem, a central result in the geometry of numbers:

**Theorem 2.1** (Siegel's formula, [24, 16]). *Let  $f \in L^1(\mathbb{R}^d, \lambda)$  where  $\lambda$  is the Lebesgue measure on  $\mathbb{R}^d$ . Define a function  $\hat{f}$  on  $X_d$  by*

$$\hat{f}(\Lambda) := \sum_{\mathbf{x} \in \Lambda \setminus \{0\}} f(\mathbf{x}).$$

*Then*

$$(2.1) \quad \int_{X_d} \hat{f} d\mu = \int_{\mathbb{R}^d} f d\lambda.$$

Note that, if two functions differ in value on a null set, (2.1) still holds, so there is no need to distinguish between functions that differ on null sets.

The space of affine unimodular lattices  $Y_d$  can be identified with the space

$$\mathrm{SL}(d, \mathbb{R}) \ltimes \mathbb{R}^d / \mathrm{SL}(d, \mathbb{Z}) \ltimes \mathbb{Z}^d.$$

In other words, it is a fiber bundle over  $X_d$  with (compact) fiber over a unimodular lattice  $\Lambda$  given by the torus  $\mathbb{R}^d / \Lambda$ . It has a natural probability measure  $\nu = \nu_d$ . As above, given  $f \in L^1(\mathbb{R}^d, \lambda)$ , define a function  $\tilde{f}$  on  $Y_d$  by

$$\tilde{f}(\Lambda + \mathbf{v}) := \sum_{\mathbf{x} \in \Lambda + \mathbf{v}} f(\mathbf{x}).$$

As above, we have

$$\int_{Y_d} \tilde{f} d\nu = \int_{\mathbb{R}^d} f d\lambda,$$

which is Siegel's formula for affine unimodular lattices; for a proof, see [18, Corollary 5.2].

**2.1.1. Primitive Vectors.** There is also a version of Siegel formula for the transform associated to summing over the set of *primitive* vectors (vectors which are not non-trivial multiples of other vectors in the lattice or, equivalently, are visible from the origin), which has an extra factor  $\frac{1}{\zeta(d)}$  on the right hand side (this is the proportion of primitive vectors), and we can obtain analogous results to Theorem 1.5 for primitive vectors with a corresponding factor of  $1/\zeta(d)$ .

**2.2. Diagonal Flows.** Given  $m, n \geq 1$ , let

$$g_t := \begin{pmatrix} e^{\frac{n}{m}t} I_m & 0 \\ 0 & e^{-t} I_n \end{pmatrix}.$$

The one-parameter group  $\{g_t\}_{t \in \mathbb{R}}$  acts on the spaces  $X_d$  and  $Y_d$  by left multiplication (equivalently, by the linear action on the lattices viewed as subsets of  $\mathbb{R}^d$ ). Given  $b > 0$ , let

$$f_b$$

be the indicator function of the set

$$\mathcal{R}_b := \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^d : \|\mathbf{x}\|^m \|\mathbf{y}\|^n \leq b, 1 \leq \|\mathbf{y}\| < 2\}.$$

Note that

$$g_{-\log 2} \mathcal{R}_b = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^d : \|\mathbf{x}\|^m \|\mathbf{y}\|^n \leq b, 2 \leq \|\mathbf{y}\| < 4\};$$

again, we are thinking of these elements of  $\mathbb{R}^d$  as column vectors. Let  $\Lambda$  be a unimodular lattice and  $\mathbf{v}$  be an element in  $\mathbb{R}^d$ . The key observation in the proof of Theorem 1.5 is

$$(2.2) \quad R_{b, 2^k}(\Lambda) = \sum_{i=0}^{k-1} \widehat{f}_b(g_{\log 2}^i \Lambda)$$

and in Theorem 1.6 is

$$(2.3) \quad R_{b, 2^k}(\Lambda + \mathbf{v}) = \sum_{i=0}^{k-1} \tilde{f}_b(g_{\log 2}^i(\Lambda + \mathbf{v})).$$

**2.3. Ergodicity.** The key fact that we use is the Moore ergodicity theorem and its generalization to the space of affine unimodular lattices:

*Theorem* (Moore ergodicity [22]). The action of  $\{g_t\}$  on  $X_d$  (and  $Y_d$ ) is ergodic with respect to the Haar measure. In particular, the transformation  $g_{\log 2}$  is ergodic.

For a proof in the case of  $Y_d$ , see [13, Lemma 4.2]. Note that, by Siegel's formula above,  $\widehat{f} \in L^1(X_d, \mu)$  and  $\tilde{f} \in L^1(Y_d, \nu)$ , which allow us to apply the Birkhoff Ergodic Theorem (see, for example, Walters [30]):

*Theorem* (Birkhoff ergodic theorem). Let  $T$  be an ergodic measure-preserving transformation of a probability space  $(X, \mu)$ , and let  $f \in L^1(X, \mu)$ . Then, for almost every  $x \in X$ ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} f(T^i x) = \int_X f d\mu.$$

**2.4. Proof of Theorems 1.5 and 1.6.** Applying the Birkhoff ergodic theorem to the expressions (2.2) and (2.3) and using Siegel's formula and our volume computation from Section 1.2, we obtain, for almost every  $\Lambda \in X_d$ ,

$$(2.4) \quad \lim_{k \rightarrow \infty} \frac{1}{k} R_{b,2^k}(\Lambda) = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} \widehat{f}_b(g_{\log 2}^i \Lambda) = \int_{X_d} \widehat{f}_b d\mu = \lambda(\mathcal{R}_b) = bB_m C_n \log 2$$

and, for almost every  $(\Lambda + \mathbf{v}) \in Y_d$ ,

$$(2.5) \quad \lim_{k \rightarrow \infty} \frac{1}{k} R_{b,2^k}(\Lambda + \mathbf{v}) = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} \tilde{f}_b(g_{\log 2}^i(\Lambda + \mathbf{v})) = \int_{Y_d} \tilde{f}_b d\nu = \lambda(\mathcal{R}_b) = bB_m C_n \log 2.$$

Note that, if  $F : [0, \infty) \rightarrow [0, \infty)$  is an increasing function and  $\frac{F(2^k)}{k} \rightarrow \log 2$ , then

$$(2.6) \quad \lim_{T \rightarrow \infty} \frac{F(T)}{\log T} = 1,$$

which applied to (2.4) and (2.5) yields Theorem 1.5 and Theorem 1.6 respectively.  $\square$

**2.5. Applications.** In this section, we show how to apply the above arguments to obtain Theorems 1.3 and 1.4. (Note that Theorem 1.2 follows immediately from Theorem 1.3.) The main technique here is an approximation argument which may have further applicability to other questions arising from the interface of dynamics on the space of unimodular lattices and Diophantine approximation.

**2.5.1. Proof of Theorem 1.3.** Let

$$\Gamma := \mathrm{SL}(d, \mathbb{Z}).$$

Recall that a unimodular lattice  $\Lambda \subset \mathbb{R}^d$  can be written as  $\Lambda = g\mathbb{Z}^d$ , where  $g \in \mathrm{SL}(d, \mathbb{R})$  is well-defined up to multiplication on the right by  $\mathrm{SL}(d, \mathbb{Z})$ . Consequently,  $\Lambda$  can be uniquely expressed as

$$g\Gamma$$

when it is regarded as an element in  $X_d$ . Given a matrix  $A \in M_{m \times n}(\mathbb{R})$ , form the associated matrix

$$h_A := \begin{pmatrix} I_m & -A \\ 0 & I_n \end{pmatrix},$$

where  $I_m$  and  $I_n$  denote the  $m \times m$  and  $n \times n$  identity matrices respectively. Setting  $\Lambda_A := h_A \mathbb{Z}^d$ , a direct calculation shows that we have, for  $\tilde{b} > 0$  and  $T > 1$ ,

$$(2.7) \quad N(A, \tilde{b}, T) = R_{\tilde{b}m, T}(\Lambda_A),$$

which still holds when  $\Lambda_A$  is replaced with  $h_A \gamma \mathbb{Z}^d$  for any  $\gamma \in \Gamma$ . We also have that, as  $t \rightarrow +\infty$ ,

$$g_{-t} h_A g_t \rightarrow I_d.$$

In fact, the set

$$\mathcal{N} := \{h_A : A \in M_{m \times n}(\mathbb{R})\}$$

forms the *expanding horospherical* subgroup for  $\{g_t\}_{t \geq 0}$ , and it has a Haar measure, which we denote by  $d\lambda$ .

Recall the definition of  $\mu := \mu_d$ . Define the open set

$$\mathfrak{M} := \left\{ \begin{pmatrix} B & A \\ C & D \end{pmatrix} \in \mathrm{SL}(d, \mathbb{R}) : B \in \mathrm{GL}(m, \mathbb{R}), A \in M_{m \times n}(\mathbb{R}), C \in M_{n \times m}(\mathbb{R}), D \in M_{n \times n}(\mathbb{R}) \right\}$$

and we note that the set  $\mathrm{SL}(d, \mathbb{R}) \setminus \mathfrak{M}$  is  $\mu$ -null. Restricting the set  $\mathfrak{M}$  to only those elements for which  $A$  equals the  $m \times n$  matrix with all zero entries, a matrix which we denote by  $0$ , we obtain

$$\mathcal{H} := \left\{ \begin{pmatrix} B & 0 \\ C & D \end{pmatrix} \in \mathrm{SL}(d, \mathbb{R}) : B \in \mathrm{GL}(m, \mathbb{R}), C \in M_{n \times m}(\mathbb{R}), D \in \mathrm{GL}(n, \mathbb{R}) \right\},$$

which, by direct computation, is a subgroup of  $\mathrm{SL}(d, \mathbb{R})$ . Let  $d\mu_{\mathcal{H}}$  denote a left Haar measure on  $\mathcal{H}$ .

We now give a parametrization of  $\mathfrak{M}$ , which is, essentially, the parametrization given by J. Marklof in [17, Section 3].

**Lemma 2.2.** *Any element of  $\mathfrak{M}$  can be uniquely expressed as*

$$\begin{pmatrix} B & 0 \\ C & D \end{pmatrix} h_A$$

for some  $B \in \mathrm{GL}(m, \mathbb{R})$ ,  $C \in M_{n \times m}(\mathbb{R})$ , and  $D \in \mathrm{GL}(n, \mathbb{R})$ .

*Proof.* Let

$$\begin{pmatrix} \beta & \alpha \\ \gamma & \delta \end{pmatrix} \in \mathfrak{M}.$$

Then setting  $B = \beta$ ,  $C = \gamma$ ,  $A = -\beta^{-1}\alpha$ , and  $D = \delta - \gamma\beta^{-1}\alpha$  yields the decomposition. Since the determinant of

$$\begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$$

is 1,  $D$  is invertible. This shows that every element of  $\mathfrak{M}$  can be expressed as desired.

If an element of  $\mathfrak{M}$  has two decompositions, then we have

$$\begin{pmatrix} B' & 0 \\ C' & D' \end{pmatrix}^{-1} \begin{pmatrix} B & 0 \\ C & D \end{pmatrix} = h_{A'} h_{-A}.$$

Multiplying out the matrices shows that  $A = A'$ ,  $B = B'$ ,  $C = C'$ , and  $D = D'$ . This gives the desired uniqueness.  $\square$

Using the parametrization given by Lemma 2.2, we have that a Haar measure on  $\mathrm{SL}(d, \mathbb{R})$  is given by

$$d\mu_{\mathcal{H}} d\lambda,$$

which, thus, must be a constant multiple of  $d\mu$ . By normalizing  $d\mu_{\mathcal{H}}$  appropriately, we may assume that the constant is 1.

Consider the following submanifold of  $\mathfrak{M}/\Gamma \subset X_d$ :

$$\mathcal{W} := \{h_A \Gamma : A \in M_{m \times n}(\mathbb{R})\}.$$

We call a unimodular lattice  $\Lambda$  *Birkhoff generic* (for the action of  $g_{\log 2}$  with respect to the function  $\widehat{f}_b$ ) if

$$(2.8) \quad \lim_{j \rightarrow \infty} \frac{1}{j} \sum_{i=0}^{j-1} \widehat{f}_b(g_{\log 2}^i \Lambda) = \int_{X_d} \widehat{f}_b d\mu.$$

holds. By the ergodicity of  $g_{-\log 2}$  and the Birkhoff ergodic theorem, (2.8) holds for  $\mu$ -almost every  $\Lambda$ .

To obtain our desired result, we now use an approximation argument to show that almost every lattice in the submanifold  $\mathcal{W}$  is Birkhoff generic. Pick a lattice

$$\widetilde{\Lambda} := h_{\widetilde{A}} \Gamma$$

in  $\mathcal{W}$ . Let  $U_{\mathcal{H}}(I_d)$  be an small open ball in  $\mathcal{H}$  around the identity element  $I_d$  and  $U_{\mathcal{W}}(h_{\widetilde{A}})$  be a small open ball in  $\mathcal{W}$  around  $h_{\widetilde{A}}$ . Then, by Lemma 2.2,

$$(2.9) \quad U := U(\widetilde{\Lambda}) := U_{\mathcal{H}}(I_d) U_{\mathcal{W}}(h_{\widetilde{A}})$$

is a small open set in  $X_d$  containing  $\widetilde{\Lambda}$ . Recall that  $\mathrm{SL}(d, \mathbb{R})$  has a right  $\mathrm{SL}(d, \mathbb{R})$ -invariant metric, using which it follows that the right action of elements of  $\Gamma$  is by isometries. Fix a fundamental domain  $\mathcal{F}$  for the action of  $\Gamma$ . Recall that the lattice  $\Gamma$  acts properly discontinuously on  $\mathrm{SL}(d, \mathbb{R})$  and, thus, lifting  $U$  to  $\mathrm{SL}(d, \mathbb{R})$  using the canonical projection mapping, we obtain an infinite family of disjoint isometric sets. The  $\mu$ -measure of this family that lies in  $\mathcal{F}$  is finite and is equal to the  $\mu$ -measure of one of these disjoint sets.

We now show, roughly speaking, almost all lattices of almost all  $\mathcal{H}$ -translates of the submanifold  $\mathcal{W}$  is Birkhoff generic. Precisely, we have the following, which is simply an application of Fubini's theorem:

**Proposition 2.3.** *Let  $U$  be the open set defined in (2.9). For  $\mu_{\mathcal{H}}$ -almost every  $M \in U_{\mathcal{H}}(I_d)$ , there exists a subset  $V_M \subset U_{\mathcal{W}}(h_{\widetilde{A}})$  such that*

$$(2.10) \quad \int_{\mathcal{N}} \mathbb{1}_{V_M}(h_A) d\lambda(A) = \int_{\mathcal{N}} \mathbb{1}_{U_{\mathcal{W}}(h_{\widetilde{A}})}(h_A) d\lambda(A)$$

and, for every  $h_A \Gamma \in V_M$ , the lattice

$$Mh_A \Gamma$$

is Birkhoff generic.

*Proof.* By construction, the set  $U$  is a subset of  $\mathfrak{M}/\Gamma$  and we may, without loss of generality, regard  $U$  as a subset of  $\mathcal{F}$ . Since (2.8) holds for  $\mu$ -almost every element in  $X_d$ , there exists a set  $U_{bg} \subset U$  for which every element in  $U_{bg}$  is Birkhoff generic and for which

$$(2.11) \quad \mu(U_{bg}) = \mu(U).$$

We use the parametrization given in Lemma 2.2.

Let

$$U_{bg,M} := \{h_A \in U_{\mathcal{W}}(h_{\tilde{A}}) : Mh_A \in U_{bg}\}.$$

Fubini's theorem implies that

$$\mathbb{1}_{U_{bg}}(Mh_A)$$

is  $\mu_\lambda$ -integrable for  $\mu_H$ -almost every  $M \in U_{\mathcal{H}}(I_d)$ . If, for  $\mu_H$ -almost every  $M$ , (2.10) holds when we set

$$V_M := U_{bg,M},$$

then we have proved the proposition.

Otherwise, there exists a subset  $V$  of  $U_{\mathcal{H}}(I_d)$  of positive  $\mu_H$ -measure such that, for every element  $M \in V$ , we have that

$$(2.12) \quad \int_{\mathcal{N}} \mathbb{1}_{U_{bg}}(Mh_A) \, d\lambda(A) = \int_{\mathcal{N}} \mathbb{1}_{U_{bg,M}}(h_A) \, d\lambda(A) < \int_{\mathcal{N}} \mathbb{1}_{U_{\mathcal{W}}(h_{\tilde{A}})}(h_A) \, d\lambda(A).$$

Consequently, we have that

$$\begin{aligned} \mu(U_{bg}) &= \int_{U_{\mathcal{H}}(I_d)} \int_{U_{\mathcal{W}}(h_{\tilde{A}})} \mathbb{1}_{U_{bg}}(Mh_A) \, d\mu_{\mathcal{H}}(M) \, d\lambda(A) \\ &= \int_{U_{\mathcal{H}}(I_d) \setminus V} \int_{U_{\mathcal{W}}(h_{\tilde{A}})} \mathbb{1}_{U_{bg}}(Mh_A) \, d\mu_{\mathcal{H}}(M) \, d\lambda(A) + \int_V \int_{U_{\mathcal{W}}(h_{\tilde{A}})} \mathbb{1}_{U_{bg}}(Mh_A) \, d\mu_{\mathcal{H}}(M) \, d\lambda(A) \\ &= \int_{U_{\mathcal{H}}(I_d) \setminus V} \int_{U_{\mathcal{W}}(h_{\tilde{A}})} \mathbb{1}_{U_{\mathcal{H}}(I_d) \setminus V}(M) \mathbb{1}_{U_{bg,M}}(h_A) \, d\mu_{\mathcal{H}}(M) \, d\lambda(A) \\ &\quad + \int_V \int_{U_{\mathcal{W}}(h_{\tilde{A}})} \mathbb{1}_V(M) \mathbb{1}_{U_{bg,M}}(h_A) \, d\mu_{\mathcal{H}}(M) \, d\lambda(A) \\ &< \int_{U_{\mathcal{H}}(I_d)} \int_{U_{\mathcal{W}}(h_{\tilde{A}})} \mathbb{1}_{U_{\mathcal{H}}(I_d)}(M) \mathbb{1}_{U_{\mathcal{W}}(h_{\tilde{A}})}(h_A) \, d\mu_{\mathcal{H}}(M) \, d\lambda(A) = \mu(U), \end{aligned}$$

where we have applied (2.12) to obtain the strict inequality. This contradicts (2.11). Consequently, the set  $V$  cannot exist and the proof of the proposition is complete.  $\square$

Let  $0 < \varepsilon'_\ell \rightarrow 0$  be a given decreasing sequence indexed by  $\ell \in \mathbb{N}$ . From  $U_{\mathcal{H}}(I_d)$ , pick a sequence of elements

$$\left\{ \begin{pmatrix} B_\ell & 0 \\ C_\ell & D_\ell \end{pmatrix} \right\}_\ell$$

such that, for every element  $h_A \Gamma$  of a subset  $L_\infty \subset U_{\mathcal{W}}(h_{\tilde{A}})$ , the lattice

$$(2.13) \quad \begin{pmatrix} B_\ell & 0 \\ C_\ell & D_\ell \end{pmatrix} h_A \Gamma$$

is Birkhoff generic with respect to  $\widehat{f}_b$ ,  $\widehat{f}_{b+\varepsilon'_\ell}$ , and  $\widehat{f}_{b-\varepsilon'_\ell}$  for all  $\ell$  and such that

$$(2.14) \quad \begin{aligned} B_\ell &\xrightarrow{\ell \rightarrow \infty} I_m \\ C_\ell &\xrightarrow{\ell \rightarrow \infty} 0 \in M_{m,n}(\mathbb{R}) \\ D_\ell &\xrightarrow{\ell \rightarrow \infty} I_n. \end{aligned}$$



By Proposition 2.3, the set  $L_\infty$  is a  $\lambda$ -full measure subset of  $U_{\mathcal{W}}(h_{\bar{A}})$ . Hence, picking any lattice  $h_A\Gamma \in L_\infty$  implies that the lattices

$$\Lambda_\ell := \begin{pmatrix} B_\ell & 0 \\ C_\ell & D_\ell \end{pmatrix} h_A\Gamma$$

are Birkhoff generic with respect to  $\widehat{f}_b$ ,  $\widehat{f}_{b+\varepsilon'_\ell}$ , and  $\widehat{f}_{b-\varepsilon'_\ell}$  for all  $\ell$ .

We now approximate lattices  $h_A\Gamma \in L_\infty$  in the full measure subset  $L_\infty$  with the lattices  $\Lambda_\ell$  as follows. Let  $\gamma \in \Gamma$ . Recall that the unimodular lattice  $g\Gamma$  can be expressed as  $g\gamma\mathbb{Z}^d \subset \mathbb{R}^d$ . First consider the case of  $\gamma = I_d$ , namely the following:

$$(2.15) \quad h_A\mathbb{Z}^d, \quad \Lambda_\ell = \begin{pmatrix} B_\ell & 0 \\ C_\ell & D_\ell \end{pmatrix} h_A\mathbb{Z}^d.$$

Counting lattice points of  $h_A\mathbb{Z}^d$  in a thinning region

$$\widetilde{\mathcal{R}}_b(\mathbb{R}^d) := \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^d : \|\mathbf{x}\|^m \|\mathbf{y}\|^n \leq b, 1 \leq \|\mathbf{y}\|\},$$

namely finding the cardinality of

$$h_A\mathbb{Z}^d \cap \widetilde{\mathcal{R}}_b(\mathbb{R}^d) = \widetilde{\mathcal{R}}_b(h_A\mathbb{Z}^d),$$

is the same as counting lattice points of  $\Lambda_\ell$  in the skewed thinning region

$$\left\{ \begin{pmatrix} B_\ell & 0 \\ C_\ell & D_\ell \end{pmatrix} (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^d : \|\mathbf{x}\|^m \|\mathbf{y}\|^n \leq b, 1 \leq \|\mathbf{y}\| \right\}$$

where recall we have used the notation  $(\mathbf{x}, \mathbf{y})$  to denote a  $d$ -column vector.

Since (2.14) holds for  $\Lambda_\ell$ , we have that

$$\begin{aligned} (1 - \varepsilon_\ell) \|\mathbf{y}\| - \widetilde{\varepsilon}_\ell &\leq \|C_\ell \mathbf{x} + D_\ell \mathbf{y}\| \leq (1 + \varepsilon_\ell) \|\mathbf{y}\| + \widetilde{\varepsilon}_\ell \\ (1 - \widehat{\varepsilon}_\ell) \|\mathbf{x}\|^m \|\mathbf{y}\|^n &\leq \|B_\ell \mathbf{x}\|^m \|C_\ell \mathbf{x} + D_\ell \mathbf{y}\|^n \leq (1 + \widehat{\varepsilon}_\ell) \|\mathbf{x}\|^m \|\mathbf{y}\|^n \end{aligned}$$

where  $\varepsilon_\ell, \widetilde{\varepsilon}_\ell, \widehat{\varepsilon}_\ell \rightarrow 0$  as  $\ell \rightarrow \infty$ .

Since we have  $\|\mathbf{x}\|^m \|\mathbf{y}\|^n \leq b$  for the skewed thinning region, we have shown that we may approximate it with inner and outer thinning regions  $\widetilde{\mathcal{R}}_{b-\varepsilon'_\ell}(\mathbb{R}^d)$  and  $\widetilde{\mathcal{R}}_{b+\varepsilon'_\ell}(\mathbb{R}^d)$  up to a small precompact set. The precompact set that we might need to exclude arises as follows. The inner and outer thinning regions satisfy the constraint that  $\|\mathbf{y}\| \geq 1$ , but, for the skewed thinning region, the infimum of  $\|C_\ell \mathbf{x} + D_\ell \mathbf{y}\|$  is close to 1 but not necessarily 1. This follows because  $\|C_\ell \mathbf{x}\| \leq \widetilde{\varepsilon}_\ell$ ,  $\|D_\ell \mathbf{y}\| \geq (1 - \varepsilon_\ell) \|\mathbf{y}\|$ , and, for the skewed thinning region,  $\|\mathbf{y}\| \geq 1$ . If the infimum of  $\|C_\ell \mathbf{x} + D_\ell \mathbf{y}\|$  is less than 1, then we remove the small precompact set from the skewed thinning region so that infimum of  $\|C_\ell \mathbf{x} + D_\ell \mathbf{y}\|$  equals 1. Therefore, up to possibly removing a small precompact set, we have the following inclusions

$$\widetilde{\mathcal{R}}_{b-\varepsilon'_\ell}(\mathbb{R}^d) \subset \begin{pmatrix} B_\ell & 0 \\ C_\ell & D_\ell \end{pmatrix} \widetilde{\mathcal{R}}_b(\mathbb{R}^d) \subset \widetilde{\mathcal{R}}_{b+\varepsilon'_\ell}(\mathbb{R}^d)$$

and thus also

$$\Lambda_\ell \cap \widetilde{\mathcal{R}}_{b-\varepsilon'_\ell}(\mathbb{R}^d) \subset \begin{pmatrix} B_\ell & 0 \\ C_\ell & D_\ell \end{pmatrix} (h_A\mathbb{Z}^d \cap \widetilde{\mathcal{R}}_b(\mathbb{R}^d)) \subset \Lambda_\ell \cap \widetilde{\mathcal{R}}_{b+\varepsilon'_\ell}(\mathbb{R}^d).$$

Now let us repeat the above, namely counting lattice points, with the thinning region  $\widetilde{\mathcal{R}}_b(\mathbb{R}^d)$  replaced by its truncation

$$\widetilde{\mathcal{R}}_{b,2^j}(\mathbb{R}^d) := \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^d : \|\mathbf{x}\|^m \|\mathbf{y}\|^n \leq b, 1 \leq \|\mathbf{y}\| < 2^j\}.$$

From this, we obtain, up to removing the small precompact set already mentioned,

$$(2.16) \quad \Lambda_\ell \cap \widetilde{\mathcal{R}}_{b-\varepsilon'_\ell,2^j}(\mathbb{R}^d) \subset \begin{pmatrix} B_\ell & 0 \\ C_\ell & D_\ell \end{pmatrix} (h_A\mathbb{Z}^d \cap \widetilde{\mathcal{R}}_{b,2^j}(\mathbb{R}^d)) \subset \Lambda_\ell \cap \widetilde{\mathcal{R}}_{b+\varepsilon'_\ell,2^{j+1}}(\mathbb{R}^d),$$

where the latter inclusion arises because, on the truncated skewed thinning region, the supremum of

$$\|C_\ell \mathbf{x} + D_\ell \mathbf{y}\| \leq (1 + \varepsilon_\ell) 2^j + \widetilde{\varepsilon}_\ell < 2^{j+1}.$$

Recall the notation and discussion in Section 2.2. Now since any compact set (and thus any precompact set) of  $\mathbb{R}^d$  can only contain a finite number of lattice points, (2.16) implies that, up to a fixed finite number of lattice points, we have that

$$R_{b-\varepsilon'_\ell,2^j}(\Lambda_\ell) \leq R_{b,2^j}(h_A\mathbb{Z}^d) \leq R_{b+\varepsilon'_\ell,2^{j+1}}(\Lambda_\ell).$$

Consequently, we have that

$$\lim_{j \rightarrow \infty} \frac{1}{j} R_{b-\varepsilon'_\ell, 2^j}(\Lambda_\ell) \leq \lim_{j \rightarrow \infty} \frac{1}{j} R_{b, 2^j}(h_A \mathbb{Z}^d) \leq \lim_{j \rightarrow \infty} \frac{1}{j+1} R_{b+\varepsilon'_\ell, 2^{j+1}}(\Lambda_\ell),$$

the analog of which also holds when  $\gamma = I_d$  is replaced by any  $\gamma \in \Gamma$  in (2.15). Since  $\Lambda_\ell \in X_d$  is Birkhoff generic, we immediately have that

$$\int_{X_d} \widehat{f}_{b-\varepsilon'_\ell} d\mu \leq \lim_{j \rightarrow \infty} \frac{1}{j} R_{b, 2^j}(h_A \Gamma) \leq \int_{X_d} \widehat{f}_{b+\varepsilon'_\ell} d\mu.$$

Applying Siegel's formula and our volume computation from Section 1.2 and letting  $\ell \rightarrow \infty$  show that every lattice in  $L_\infty$  is Birkhoff generic with respect to  $\widehat{f}_b$ . As the chosen lattice  $\widetilde{\Lambda}$  and small open set  $U$  are arbitrary, this proof is valid everywhere, showing that almost every lattice in  $\mathcal{W}$  is Birkhoff generic with respect to  $\widehat{f}_b$ . Applying this to the proof in Section 2.4 and using (2.7) with  $b := \widetilde{b}^m$  proves Theorem 1.3. Theorem 1.2 is a special case of Theorem 1.3 with  $m = 1, n = k$ .  $\square$

*Remark 2.4.* As a remark, which is not necessary for our proof, we note that Lemma 2.2 implies that

$$\{MW\}_{M \in \mathcal{H}}$$

forms a smooth foliation of the open set of full measure  $\mathfrak{M}/\Gamma$  and, thus essentially, of  $X_d$  also. The leaves are  $\mathcal{H}$ -translates of the submanifold  $\mathcal{W}$ .

**2.5.2. Proof of Theorem 1.4.** This is the affine lattice case and will follow by using the argument in Section 2.5.1 with a few minor changes. The changes are as follows. An affine unimodular lattice  $\Lambda + \mathbf{v}$  is uniquely determined by a unimodular lattice  $\Lambda$  and a vector  $\mathbf{v} \in \mathbb{R}^d/\Lambda$ . Using Lemma 2.2 and the fact that  $\mathcal{H}$  is a group, we have that any affine unimodular lattice  $\Lambda + \mathbf{v}$  with  $\Lambda \in \mathfrak{M}/\Gamma$  can be uniquely written as

$$\begin{pmatrix} B & 0 \\ C & D \end{pmatrix} \left( h_A \Gamma + \begin{pmatrix} \mathbf{v}'_1 \\ \mathbf{v}'_2 \end{pmatrix} \right)$$

where

$$\begin{pmatrix} \mathbf{v}'_1 \\ \mathbf{v}'_2 \end{pmatrix} \in \mathbb{R}^d / (h_A \Gamma \mathbb{Z}^d).$$

We note such affine unimodular lattices form an open set whose complement has zero Haar measure. Also note that the parametrization for the affine lattice case is given by

$$\left( \mathcal{H} \ltimes \begin{pmatrix} \mathbf{0} \\ \mathbb{R}^n \end{pmatrix} \right) \left( \mathcal{N} \ltimes \begin{pmatrix} \mathbb{R}^m \\ \mathbf{0} \end{pmatrix} \right).$$

Given a pair  $(A, \mathbf{w}) \in M_{m \times n}(\mathbb{R}) \times \mathbb{R}^m$ , form the associated affine unimodular lattice  $\Lambda_A + \bar{\mathbf{w}} = h_A \gamma \mathbb{Z}^d + \bar{\mathbf{w}}$ , where  $\bar{\mathbf{w}}$  is the vector  $(\mathbf{w}, \mathbf{0}) \in \mathbb{R}^d / (h_A \gamma \mathbb{Z}^d)$  and  $\gamma \in \Gamma$ . As above, a direct calculation shows that we have, for  $\widetilde{b} > 0$  and  $T > 1$ ,

$$N(A, \mathbf{w}, \widetilde{b}, T) = R_{\widetilde{b}^m, T}(\Lambda_A + \bar{\mathbf{w}})$$

for every  $\gamma \in \Gamma$ .

The analog of  $\mathcal{W}$  from Section 2.5.1 is

$$\mathcal{W}_{\text{aff}} := \left\{ h_{\widetilde{A}} \Gamma + \begin{pmatrix} \widetilde{\mathbf{w}} \\ \mathbf{0} \end{pmatrix} : \widetilde{A} \in M_{m \times n}(\mathbb{R}) \text{ and } \widetilde{\mathbf{w}} \in \mathbb{R}^m \right\}.$$

Finally, we note that a thinning region can be approximated by inner and outer thinning regions translated by

$$\begin{pmatrix} \mathbf{0} \\ \mathbf{v}_\ell \end{pmatrix} \in \mathbb{R}^m \times \mathbb{R}^n$$

where  $\mathbf{v}_\ell \rightarrow \mathbf{0}$  as  $\ell \rightarrow \infty$ . (This is because the directions extending to infinity of the thinning region is preserved by such translations.) We note that this approximation is in addition to the approximation of Section 2.5.1. This shows that almost every affine lattice in  $\mathcal{W}_{\text{aff}}$  is Birkhoff generic and, thus, proves Theorem 1.4.  $\square$

Note the analog of Remark 2.4 applies here.

### 3. TRANSLATION SURFACES

We now apply our technique to translation surfaces, giving a proof of Theorem 1.1. While above we appeal to the Siegel formula, in the setting of translation surfaces, we use the Siegel-Veech formula.

**Theorem 3.1** (Siegel-Veech formula, [11], pg. 584). *Let  $f$  be a continuous function of compact support on  $\mathbb{R}^2$ . Define  $\hat{f}$ , a function on the stratum  $\mathcal{H}$ , by*

$$\hat{f}(\omega) := \sum_{\mathbf{v} \in \Lambda_\omega} f(\mathbf{v});$$

*here  $\Lambda_\omega$  refers to the set of holonomy vectors on the translation surface  $\omega$ . Let  $\mu$  be a  $\mathrm{SL}(2, \mathbb{R})$ -invariant measure on the stratum  $\mathcal{H}$ . Then there exists  $C = C(\mu)$ .*

$$(3.1) \quad \int_{\mathcal{H}} \hat{f}(\omega) d\mu(\omega) = C \int_{\mathbb{R}^2} f d\lambda,$$

where  $\lambda$  is the Lebesgue measure on  $\mathbb{R}^2$ .

*Proof of Theorem 1.1:* Let

$$\mathcal{R}_i := \{(x, y) \in \mathbb{R}^2 : |xy| \leq b \text{ and } 2^{i-1} \leq y < 2^i\}$$

for  $i = 1, \dots$ , and define

$$g_t := \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}.$$

Noting that  $g_{-\log 2} \mathcal{R}_i = \mathcal{R}_{i+1}$ , we have that

$$R_{b,2^k}(\omega) = \sum_{i=0}^{k-1} \#(\mathcal{R}_i \cap \Lambda_\omega) = \sum_{i=0}^{k-1} \#(\mathcal{R}_1 \cap g_{\log 2}^i \Lambda_\omega) = \sum_{i=0}^{k-1} \mathbb{1}_{\mathcal{R}_1}(g_{\log 2}^i \Lambda_\omega).$$

By assumption,  $\mu$  is an ergodic invariant measure for the action of  $g_{\log 2}$  on  $\mathcal{H}$ ; by the Birkhoff ergodic theorem, the Siegel-Veech formula, and our volume computation from Section 1.2, we have

$$\lim_{k \rightarrow \infty} \frac{R_{b,2^k}(\omega)}{k} = \int_{\mathcal{H}} \hat{\mathbb{1}}_{\mathcal{R}_1} d\mu = C\lambda(\mathcal{R}_1) = 2bC \log 2.$$

Here  $C = C(\mu)$ . Note that, to apply the Siegel-Veech formula for the indicator function  $\mathbb{1}_{\mathcal{R}_1}$ , we proceed as follows. Taking an outer approximation given by a tessellation of small closed squares, using the Urysohn lemma, and using the Siegel-Veech formula for continuous functions of compact support, derive an upper bound. Similarly, for an inner approximation of small open squares, derive a lower bound. Taking the limit, the desired relation (3.1) holds. Applying (2.6) to  $F(t) := R_{b,t}(\omega)$  yields the result.  $\square$

**Acknowledgements.** We are grateful to the referees for their helpful comments.

### REFERENCES

- [1] J. S. Athreya and J. Chaika. The distribution of gaps for saddle connection directions. *Geom. Funct. Anal.*, 22(6):1491–1516, 2012.
- [2] J. S. Athreya, J. Chaika, and S. Lelievre. The gap distribution of slopes on the golden L, 2013, arXiv:1308.4203, to appear, *Contemporary Mathematics*.
- [3] J. S. Athreya, A. Ghosh, and J. Tseng. Spherical averages of Siegel transforms for higher rank diagonal actions and applications. *ArXiv e-prints*, July 2014, 1407.3573.
- [4] J. S. Athreya, A. Ghosh, and J. Tseng. Spiraling of approximations and spherical averages of siegel transforms. *Journal of the London Mathematical Society*, 91(2):383–404, 2015.
- [5] V. Bernik, D. Kleinbock, and G. A. Margulis. Khintchine-type theorems on manifolds: the convergence case for standard and multiplicative versions. *Internat. Math. Res. Notices*, (9):453–486, 2001.
- [6] J. Chaika, Y. Cheung, and H. Masur. Winning games for bounded geodesics in moduli spaces of quadratic differentials. *J. Mod. Dyn.*, 7(3):395–427, 2013.
- [7] M. Einsiedler and J. Tseng. Badly approximable systems of affine forms, fractals, and Schmidt games. *J. Reine Angew. Math.*, 660:83–97, 2011.
- [8] A. Eskin, G. Margulis, and S. Mozes. Upper bounds and asymptotics in a quantitative version of the Oppenheim conjecture. *Ann. of Math. (2)*, 147(1):93–141, 1998.
- [9] A. Eskin and H. Masur. Asymptotic formulas on flat surfaces. *Ergodic Theory Dynam. Systems*, 21(2):443–478, 2001.
- [10] A. Eskin, M. Mirzakhani, and A. Mohammadi. Isolation, equidistribution, and orbit closures for the  $\mathrm{sl}(2, \mathbb{R})$  action on moduli space, 2013, arXiv:1305.3015, to appear, *Annals of Mathematics*.
- [11] B. Hasselblatt and A. Katok, editors. *Handbook of dynamical systems. Vol. 1B*. Elsevier B. V., Amsterdam, 2006.

- [12] A. Khintchine. Zur metrischen Theorie der diophantischen Approximationen. *Math. Z.*, 24(1):706–714, 1926.
- [13] D. Kleinbock. Badly approximable systems of affine forms. *J. Number Theory*, 79(1):83–102, 1999.
- [14] D. Kleinbock and B. Weiss. Bounded geodesics in moduli space. *Int. Math. Res. Not.*, (30):1551–1560, 2004.
- [15] D. Y. Kleinbock and G. A. Margulis. Logarithm laws for flows on homogeneous spaces. *Invent. Math.*, 138(3):451–494, 1999.
- [16] A. M. Macbeath and C. A. Rogers. Siegel’s mean value theorem in the geometry of numbers. *Proc. Cambridge Philos. Soc.*, 54:139–151, 1958.
- [17] J. Marklof. The asymptotic distribution of Frobenius numbers. *Invent. Math.*, 181(1):179–207, 2010.
- [18] J. Marklof and A. Strömbergsson. Free path lengths in quasicrystals. *Comm. Math. Phys.*, 330(2):723–755, 2014.
- [19] H. Masur. Interval exchange transformations and measured foliations. *Ann. of Math. (2)*, 115(1):169–200, 1982.
- [20] H. Masur. The growth rate of trajectories of a quadratic differential. *Ergodic Theory Dynam. Systems*, 10(1):151–176, 1990.
- [21] H. Masur and S. Tabachnikov. Rational billiards and flat structures. In *Handbook of dynamical systems, Vol. 1A*, pages 1015–1089. North-Holland, Amsterdam, 2002.
- [22] C. C. Moore. Ergodicity of flows on homogeneous spaces. *Amer. J. Math.*, 88:154–178, 1966.
- [23] W. Schmidt. A metrical theorem in diophantine approximation. *Canad. J. Math.*, 12:619–631, 1960.
- [24] C. L. Siegel. A mean value theorem in geometry of numbers. *Ann. of Math. (2)*, 46:340–347, 1945.
- [25] J. Tseng. On circle rotations and the shrinking target properties. *Discrete Contin. Dyn. Syst.*, 20(4):1111–1122, 2008.
- [26] J. Tseng. Badly approximable affine forms and Schmidt games. *J. Number Theory*, 129(12):3020–3025, 2009.
- [27] C. Uyanik and G. Work. The distribution of gaps for saddle connections on the octagon, 2014, arXiv:1409.0830.
- [28] W. A. Veech. Gauss measures for transformations on the space of interval exchange maps. *Ann. of Math. (2)*, 115(1):201–242, 1982.
- [29] W. A. Veech. Siegel measures. *Ann. of Math. (2)*, 148(3):895–944, 1998.
- [30] P. Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1982.
- [31] A. Zorich. Flat surfaces. In *Frontiers in number theory, physics, and geometry. I*, pages 437–583. Springer, Berlin, 2006.

J.S.A.: DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, URBANA, IL 61801  
*E-mail address:* jathreya@illinois.edu

A.P.: DEPARTMENT OF MATHEMATICS & COMPUTER SCIENCE, EASTERN ILLINOIS UNIVERSITY, CHARLESTON, IL 61920  
*E-mail address:* ajnparrish@gmail.com

J.T.: SCHOOL OF MATHEMATICS, UNIVERSITY OF BRISTOL, BRISTOL BS8 1TW, U.K.  
*E-mail address:* j.tseng@bristol.ac.uk